BAB II

LANDASAN TEORI

Bab ini membahas pengertian-pengertian dasar yang digunakan sebagai landasan pembahasan pada bab selanjutnya. Pengertian-pengertian dasar yang di bahas adalah sebagai berikut:

A. Peluang

Definisi (Walpole: 90):

Peluang suatu kejadian A, disimbolkan dengan $P(A)$, adalah jumlah peluang semua titik contoh dalam A. Dengan demikian:

1. $0 \leq P(A) \leq 1$
2. $P(\emptyset) = 0$
3. $P(S) = 1$

Jika suatu percobaan mempunyai N hasil percobaan yang berbeda, dan masing-masing mempunyai kemungkinan yang sama untuk terjadi, dan bila tepat n diantara hasil percobaan itu menyusun kejadian A, maka peluang kejadian A adalah

$$P(A) = \frac{n}{N}$$

Dalil 4.11 (Walpole: 94):

Jika A dan A' adalah dua kejadian yang satu merupakan komplemen lainnya, maka

$$P(A) + P(A') = 1 \tag{2.1}$$

Bukti : Karena $A \cup A' = S$, dan kejadian A dan A' saling terpisah, maka

$$1 = P(S)$$

$= P(A \cup A')$

$= P(A) + P(A')$ (2.2)

B. Fungsi Distribusi Kumulatif

Menurut Walpole dan Myers (1995: 60) Fungsi distribusi kumulatif atau probabilitas kumulatif sering disebut fungsi distribusi saja. Fungsi distribusi variabel acak kontinu X yang dinotasikan $F(x) = P(X \leq x)$ untuk semua bilangan riil x, didefinisikan dengan:

$$F(x) = \int_{-\infty}^{x} f(t)dt \quad (2.3)$$

Sifat-sifat fungsi distribusi:

1. $\lim_{x \to \infty} F(x) = 1$
2. $\lim_{x \to -\infty} F(x) = 0$
3. Fungsi tersebut tidak turun, yaitu jika $b \geq a$ maka $F(b) \geq F(a)$
4. Fungsi tersebut kontinu dari kanan, yaitu untuk seluruh x dan $\delta > 0$

   $\lim_{\delta \to 0} F(x + \delta) - F(x) = 0$
5. $P(X > x) = 1 - F(x)$ atau $P(X > x) = 1 - P(X \leq x)$

C. Klasifikasi Data

Menurut Hasan (2004:19) suatu data dapat diklasifikasikan menjadi empat macam yaitu berdasarkan sumber pengambilan, waktu pengumpulan, sifat data dan tingkat pengukuran. Klasifikasi data diuraikan sebagai berikut:

1. Berdasarkan Sumber Pengambilannya

   Berdasarkan sumber pengambilannya, data dibedakan menjadi dua yaitu data primer dan data sekunder.

a) Data Primer

Data primer adalah data yang diperoleh atau dikumpulkan langsung di lapangan oleh orang yang melakukan penelitian atau yang bersangkutan yang memerlukannya. Data primer disebut juga data asli atau data baru.

Contoh: data kuesioner, data survei, data observasi dan sebagainya.

b) Data Sekunder

Data sekunder adalah data yang diperoleh atau dikumpulkan oleh orang yang melakukan penelitian dari sumber-sumber yang telah ada. Data ini biasanya diperoleh dari perpustakaan atau dari laporan-laporan penelitian terdahulu.

Contoh: data yang sudah tersedia di tempat-tempat tertentu seperti perpustakaan, BPS (Badan Pusat Statistik), kantor-kantor.

2. Berdasarkan Waktu Pengumpulannya

Berdasarkan waktu pengumpulannya, data dibedakan menjadi dua yaitu data berkala (*Time Series*) dan data *cross section*.

a) Data Berkala (*Time Series*)

Data berkala (*Time Series*) adalah data yang terkumpul dari waktu ke waktu untuk memberikan gambaran perkembangan suatu kegiatan atau keadaan.

Contoh: data perkembangan harga sembilan macam bahan pokok selama 10 bulan terakhir yang dikumpulkan setiap bulan.

b) Data *Cross Section*

Data *cross section* adalah data yang terkumpul pada suatu waktu tertentu untuk memberikan gambaran perkembangan suatu kegiatan atau keadaan pada waktu itu.

Contoh: data sensus penduduk tahun 2010.

3. Berdasarkan Sifat Data

Berdasarkan sifatnya, data dibedakan menjadi dua yaitu data kualitatif dan data kuantitatif.

a) Data Kualitatif

Data kualitatif adalah data yang tidak berbentuk bilangan.

Contoh: jenis kelamin, agama, warna.

b) Data Kuantitatif

Data kuantitatif adalah data yang berbentuk bilangan.

Contoh: tinggi, panjang, umur.

4. Berdasarkan Tingkat Pengukurannya

Berdasarkan tingkat pengukurannya (skala), data dibedakan menjadi empat yaitu data nominal, data ordinal, data interval dan data rasio.

a) Data Nominal

Data nominal adalah data yang berasal dari pengelompokan peristiwa berdasarkan kategori tertentu yang perbedaannya hanyalah menunjukkan perbedaan kualitatif.

Contoh: Jenis kelamin manusia misal 1 disimbolkan untuk pria dan 0 untuk wanita.

b) Data Ordinal

Data ordinal adalah data yang berasal dari objek atau kategori yang disusun menurut besarnya, dari tingkat terendah ke tingkat tertinggi atau sebaliknya, dengan jarak atau rentang yang tidak harus sama.

Contoh: mengubah nilai ujian ke nilai prestasi yaitu nilai dari 80 – 100 adalah A, nilai dari 65 – 79 adalah B dan seterusnya.

c) Data Interval

Data interval adalah data yang berasal dari objek atau kategori yang diurutkan berdasarkan suatu atribut tertentu, dimana jarak antara tiap kategori adalah sama. Pada data ini tidak terdapat angka nol absolut.

Contoh: Suhu

d) Data Rasio

Data rasio adalah data yang menghimpun semua ciri dari data nominal, data ordinal dan data interval. Pada data ini terdapat angka nol absolut.

Contoh: berat badan, panjang benda, jumlah satuan benda.

D. Distribusi Bernoulli

Definisi: Fungsi Peluang Bernoulli

Menurut Bain dan Engelhardt (1992: 91) sebuah eksperimen Bernoulli terpenuhi ketika eksperimen tersebut memiliki dua kemungkinan yang terjadi yaitu sukses atau gagal.

Variabel acak X dikatakan berdistribusi Bernoulli jika dan hanya jika fungsi peluangnya berbentuk:

$$p(x) = P(X = x) = p^x (1 - p)^{1-x}\ ;\ x = 0, 1$$

dengan mean $\mu = p$ dan varian $\sigma^2 = pq$

Bukti

$$\mu = E(X)$$
$$= \sum_{i=1}^{2} x_1 f(x_i)$$
$$= 1(p) + 0(1-p)$$
$$= p$$

$$E(X^2) = \sum_{i=1}^{2} x_1 f(x_i)$$
$$= x_1^2 f(x_1) + x_2^2 f(x_2)$$
$$= 1^2 p + 0^2 q$$
$$= p$$

$$\sigma^2 = E(X^2) - E(X)^2$$
$$= p - p^2$$
$$= p(1-p)^2$$
$$= pq$$

E. Model Peluang Linier

Menurut J. Scott Long (1997: 35) Model peluang linier merupakan bentuk model regresi yang diterapkan pada variabel tak bebas biner. Sehingga sering disebut juga model pilihan biner (*binary choice model*). Model regresinya adalah:

$Y_i = \beta_i X_i + \varepsilon_i$, dengan $i = 1, 2, \ldots, n$.

Dengan β_i adalah vektor parameter, X_i adalah vektor nilai untuk *i*-obsevasi, dan ε_i adalah galat. Persamaan tersebut ekuivalen dengan

$$Y_i = \beta_{1i} X_{i1} + \beta_{2i} X_{i2} + \ldots + \beta_{ki} X_{ik} + \varepsilon_i \tag{2.4}$$

Asumsi yang harus dipenuhi adalah mean dari ε_i atau $E(\varepsilon_i) = 0$ dan Y_i diasumsikan berdistribusi Bernoulli. Bentuk Persamaan (2.4) mempunyai tipe yang menyerupai model regresi linier, akan tetapi karena variabel Y_i berupa *binary choice* maka disebut model peluang linier.

Misalkan p_i adalah peluang dimana $Y_i = 1$, sehingga dari Persamaan (2.4) diperoleh

$$1 = \beta_{1i} + \beta_{2i} X_{i1} + \ldots + \beta_{ki} X_{ik} + \varepsilon_i$$

$$\varepsilon_i = 1 - (\beta_{1i} + \beta_{2i} X_{i1} + \ldots + \beta_{ki} X_{ik})$$

$$= 1 - \beta_{1i} - \beta_{2i} X_{i1} - \ldots - \beta_{ki} X_{ik} \tag{2.5}$$

Dan dimisalkan $1 - p_i$ adalah peluang dimana $Y_i = 0$, sehingga dari persamaan (2.4) diperoleh

$$0 = \beta_{1i} + \beta_{2i} X_{i1} + \ldots + \beta_{ki} X_{ik} + \varepsilon_i$$

$$\varepsilon_i = 0 - (\beta_{1i} + \beta_{2i} X_{i1} + \ldots + \beta_{ki} X_{ik})$$

$$= -\beta_{1i} - \beta_{2i} X_{i1} - \ldots - \beta_{ki} X_{ik} \tag{2.6}$$

Variabel acak ε_i yang berdistribusi Bernoulli mempunyai dua hasil yang mungkin. Sesuai dengan estimator tak bias maka nilai harapan ε_i, diasumsikan bahwa $E(\varepsilon_i)$ harus sama dengan nol, diperoleh:

$$E(\varepsilon_i) = p_i(Y_i = 1|X_i) + (1 - p_i)(Y_i = 0|X_i) = 0$$

$$= p_i(1 - \beta_{1i} - \beta_{2i} X_{i1} - \ldots - \beta_{ki} X_{ik}) + (1 - p_i)(-\beta_{1i} - \beta_{2i} X_{i1} - \ldots - \beta_{ki} X_{ik})$$

$= p_i - \beta_{1i} - \beta_{2i} X_{i1} - \ldots - \beta_{ki} X_{ik}$

atau

$$p_i = \beta_{1i} + \beta_{2i} X_{i1} + \ldots + \beta_{ki} X_{ik} \quad (2.7)$$

varian dari ε_i atau σ_i^2 adalah $E(\varepsilon_i^2)$ dan karena $E(\varepsilon_i^2)$ diasumsikan sama dengan nol, maka diperoleh:

$\mathrm{Var}(\varepsilon_i) = \sigma_i^2 = p_i(Y_i = 1|X_i)^2 + (1 - p_i)(Y_i = 0|X_i)^2$

$= p_i (1 - \beta_{1i} - \beta_{2i} X_{i1} - \ldots - \beta_{ki} X_{ik})^2 + (1 - p_i)(-\beta_{1i} - \beta_{2i} X_{i1} - \ldots - \beta_{ki} X_{ik})^2$

$= p_i (1 - p_i)^2 + (1 - p_i)\, p_i^{\,2}$

$= p_i (1 - p_i)$

$= 0$

$\mathrm{Cov}(\varepsilon_i, \varepsilon_j) = E[(\varepsilon_i - E(\varepsilon_i))(\varepsilon_j - E(\varepsilon_j))]$, dengan $i \neq j$

$= E[\varepsilon_i\, \varepsilon_j - E(\varepsilon_i)\, \varepsilon_j - E(\varepsilon_j)\, \varepsilon_i + E(\varepsilon_i) E(\varepsilon_j)]$

$= 0$, karena ε_i dan ε_j independen

Dari pernyataan diatas p_i adalah peluang $Y_i = 1$(kejadian terjadi) dan $1 - p_i$ adalah peluang $Y_i = 0$ (kejadian tidak terjadi). Karena Y_i hanya memiliki dua kejadian yang mungkin terjadi maka Y_i juga mengikuti distribusi Bernoulli seperti ε_i.

Distribusi Bernoulli mempunyai mean p dan varian $p(1 - p)$, sehingga diperoleh:

$E(Y_i) = Y_i\,(p(Y_i = 1|X_i)) + Y_i\,(p(Y_i = 0|X_i))$

$= 1(p_i) + 0(1 - p_i)$

$$= p_i \quad (2.8)$$

$E(Y_i|X_i) = p(Y_i = 1)(Y_i = 1|X_i) + p(Y_i = 0)(Y_i = 0|X_i)$

$= p_i(1 - \beta_{1i} - \beta_{2i} X_{i1} - \ldots - \beta_k X_{ik}) + (1 - p_i)(-\beta_{1i} - \beta_{2i} X_{i1} - \ldots - \beta_{ki} X_{ik})$

$$= p_i - \beta_{1i} - \beta_{2i}X_{i1} - \ldots - \beta_{ki}X_{ik} \quad (2.9)$$

Persamaan (2.9) adalah nilai harapan bersyarat dari Persamaan (2.4) dan dapat dinyatakan sebagai peluang bersyarat dari Y_i.

Karena peluang p_i harus terletak pada interval 0 dan 1 maka batasan $E(Y_i|X_i)$ adalah $0 \leq E(Y_i|X_i) \leq 1$. Sehingga dapat dikatakan bahwa nilai harapan bersyarat terletak pada interval 0 dan 1.

F. Model Variabel Laten

Menurut J. Scott Long (1997: 40) Model variabel laten biasanya digunakan ketika asumsi-asumsi dalam model pilihan biner tidak dibuat. Artinya asumsi dari variabel tak bebas Y_i tidak diketahui. Misalkan terdapat pilihan dari wanita yang sudah menikah bekerja atau tidak. Perbedaan antara bekerja atau tidak terletak pada berapa banyak gaji dan karakteristik seseorang, seperti usia, pendidikan, mempunyai anak atau belum, dan lain-lain. Sehingga perbedaan dalam Y_i antara bekerja atau tidak merupakan fungsi dari berbagai macam karakteristik yang diamati sebagai X_i dan karakteristik yang tidak diamati sebagai ε_i.

Model regresi untuk variabel laten adalah:

$$Y_i^* = \beta_1 + \beta_2X_{i1} + \ldots + \beta_kX_{ik} + \varepsilon_i = \beta' X_i + \varepsilon_i \quad (2.10)$$

Karena Y_i^* merupakan variabel laten, maka yang diamati dari Y_i^* adalah keadaan dimana $Y_i = 1$ jika dan hanya jika $Y_i^* > 0$ dan $Y_i = 0$ untuk yang lain, maka diperoleh:

$$P(Y_i = 1) = P(Y_i^* > 0)$$

$$= P(\beta' X_i + \varepsilon_i > 0)$$

$= P(\varepsilon_i > -\beta' X_i)$

$$= P\left(\frac{\varepsilon_i}{\sigma} > -\frac{\beta' X_i}{\sigma}\right)$$

$$= 1 - F(-\beta' X_i)$$

$$= F(\beta' X_i) \quad (2.11)$$

Dimana F menyatakan fungsi distribusi dari ε_i. Jika dipilih distribusi normal standar maka akan terbentuk model probit dengan asumsi $\varepsilon_i \sim N(0,1)$ dan ε_i bebas untuk semua X_i, yaitu

$$Y_i^* = \beta' X_i + \varepsilon_i$$

Atau ekuivalen dengan

$$Y_i = 1 \text{ jika } Y_i^* > 0 \text{ dan } Y_i = 0 \text{ jika } Y_i^* \leq 0$$

G. Metode Maksimum Likelihood

Menurut Bain dan Engelhardt (1992: 293) Metode maksimum likelihood merupakan salah satu cara untuk melakukan penaksiran parameter yang tidak diketahui. Prosedur penaksiran maksimum likelihood menguji apakah penaksiran maksimum yang tidak diketahui dari fungsi likelihood suatu sampel nilainya sudah memaksimumkan fungsi likelihood.

Misalkan $X_1, X_2, \ldots, X_n$ adalah variabel acak dari populasi dengan fungsi densitas peluangnya dinyatakan oleh $f(x, \theta)$, dengan θ adalah parameter yang tidak diketahui. Maka fungsi likelihood sampel tersebut adalah:

$$L(x_1, x_2, \ldots, x_n; \theta) = f(x_1; \theta) f(x_2; \theta) \ldots f(x_n; \theta)$$

$$= \prod_{i=1}^{n} f(x_i, \theta)$$

$$= L(\theta | x_1, x_2, \ldots, x_n)$$

$$= L(\theta) \tag{2.12}$$

Kemudian Persamaan (2.12) tersebut didiferensialkan terhadap θ untuk memperoleh penaksiran yang maksimum.

Dalam banyak kasus, penggunaan diferensiasi akan lebih mudah bekerja pada logaritma natural dari $L(x_1, x_2, \dots, x_n; \theta)$, yaitu:

$$\ln L(x_1, x_2, \dots, x_n; \theta) \tag{2.14}$$

Langkah-langkah untuk menentukan penaksiran maksimum likelihood dari θ_i adalah:

1. Menentukan fungsi likelihood

   $L(x_1, x_2, \dots, x_n; \theta) = f(x_1, \theta) f(x_2, \theta) \dots f(x_n, \theta),$

2. Membentuk logaritma natural likelihood

   $\ln L(x_1, x_2, \dots, x_n; \theta) = \ln (f(x_1, \theta) f(x_2, \theta) \dots f(x_n, \theta))$

3. Menurunkan persamaan logaritma natural likelihood terhadap θ dan menyelesaikannya

   $$\frac{\partial \ln L(x_1, x_2, \dots, x_n; \theta)}{\partial \theta} = 0$$

4. Didapat penaksiran maksimum likelihood θ

Contoh:

Tentukan estimator maksimum likelihood (MLE) untuk θ berdasarkan sampel acak berukuran n dari fungsi $f(x; \theta) = \theta x^{\theta - 1}$; $0 < x < 1$; $0 < \theta$

Jawab

Dari soal tersebut dapat ditentukan fungsi likelihoodnya sebagai berikut:

$$f(x; \theta) = \theta x^{\theta - 1}$$

$$L(x_1, x_2, \dots, x_n; \theta) = \theta x_1^{\theta - 1} \theta x_2^{\theta - 1} \dots \theta x_n^{\theta - 1}$$

$$= \theta^n (x_1 . x_2 . \dots . x_n)^{\theta - 1}$$

Setelah fungsi likelihoodnya didapat, langkah selanjutnya adalah membentuk logaritma likelihood dari fungsi tersebut. Berikut adalah bentuk logaritma natural dari fungsi likelihood $\theta^n (x_1 . x_2 . \dots . x_n)^{\theta - 1}$

$$\ln L(x_1, x_2, \dots, x_n ; \theta) = \ln \theta x_1^{\theta - 1} \theta x_2^{\theta - 1} \dots \theta x_n^{\theta - 1}$$

$$= \ln \theta^n (x_1 . x_2 . \dots . x_n)^{\theta - 1}$$

$$= \ln \theta^n + \ln (x_1 . x_2 . \dots . x_n)^{\theta - 1}$$

$$= n \ln \theta + (\theta - 1) \ln (x_1 . x_2 . \dots . x_n)$$

Untuk memperoleh nilai penaksiran yang maksimum maka dari fungsi logaritma natural likelihood yang diperoleh diturunkan terhadap θ.

$$\frac{\partial \ln L(\theta)}{\partial \theta} = 0$$

$$\frac{\partial (n \ln \theta + (\theta - 1) \ln (x_1 . x_2 . \dots . x_n))}{\partial \theta} = 0$$

$$\frac{\mathrm{n}}{\theta} + \ln (x_1 . x_2 . \dots . x_n) = 0$$

$$\frac{\mathrm{n}}{\theta} = -\ln (x_1 . x_2 . \dots . x_n)$$

$$\frac{\theta}{\mathrm{n}} = -\frac{1}{\ln (x_1 . x_2 . \dots . x_n)}$$

$$\frac{\theta}{\mathrm{n}} = -\frac{1}{\sum \ln X_i}$$

$$\theta = -\frac{\mathrm{n}}{\sum \ln X_i}$$

H. Model Regresi Probit

Menurut Greene (2003: 669) Model regresi probit adalah model linear $Y_i^* = \beta' X_i + \varepsilon_i$ yang menggunakan bilangan biner atau variabel *dummy* sebagai variabel tak bebasnya dan mengandaikan galat ε_i berdistribusi normal $N(0, \sigma^2)$. Variabel *dummy* yang dimaksud disini adalah jenis variabel diskret yang mempunyai dua nilai.

Misalkan terdapat variabel Y_i^* yang menunjukkan sentimen atau perasaan individu terhadap suatu hal, contohnya sikap seseorang terhadap suatu partai politik tertentu. Sikap tersebut digunakan sebagai variabel tak bebas dan variabel tak bebas ini dipengaruhi oleh berbagai karakteristik individu dan kondisi lingkungan, sebagai variabel bebasnya, sehingga persamaan Y_i^* dapat dituliskan sebagai:

$$Y_i^* = \beta' X_i + \varepsilon_i \tag{2.15}$$

dengan β' adalah faktor koefisien, X_i adalah faktor peubah bebas, dan ε_i adalah faktor galat yang diasumsikan berdistribusi normal.

Y_i^* tidak bisa diamati, tetapi tindakan atau pilihan tindakan individu tersebut bisa diamati jika Y_i^* melewati batas tertentu. Misalnya jika $Y_i^* > 0$, maka $Y_i = 1$ dan jika $Y_i^* \leq 0$, maka $Y_i = 0$. Dari hal tersebut diperoleh

$$P(Y_i = 1) = P(Y_i^* > 0) = P(\beta' X_i + \varepsilon_i > 0)$$

$$= P(\varepsilon_i > -\beta' X_i)$$

$$= P\left(\frac{\varepsilon_i}{\sigma} > -\frac{\beta' X_i}{\sigma}\right) \tag{2.16}$$

$$= 1 - F(-\beta' X_i)$$

$$= F(\beta' X_i) \tag{2.17}$$

Maka dari persamaan (2.17) diperoleh

$$P(Y_i = 0) = 1 - F(\beta' X_i) \tag{2.18}$$

Model dengan peluang sukses $F(\beta' X_i)$ dan peluang gagal $1 - F(\beta' X_i)$ dari pengamatan n yang saling bebas sesuai distribusi Bernoulli fungsi likelihoodnya adalah perkalian dari peluang tiap observasinya.

$$L = P(Y_1 = \mathrm{y}_1 . Y_2 = \mathrm{y}_2 \ldots . Y_\mathrm{n} = \mathrm{y}_\mathrm{n})$$

$$= \left(1 - F(\beta' X_1)\right)\left(1 - F(\beta' X_2)\right) \ldots . \left(1 - F(\beta' X_n)\right)$$

$$\times F(\beta' X_1)\, F(\beta' X_2) \ldots F(\beta' X_n)$$

$$= \prod_{y_i=0}^{n} \left(1 - F\left(\beta' X_i\right)\right) \prod_{y_i=1}^{\mathrm{n}} F\left(\beta' X_i\right)$$

$$= \prod_{i=1}^{n} F(\beta' X_i)^{y_i} \left(1 - F(\beta' X_i)\right)^{1-y_i} \tag{2.19}$$

Dengan melakukan logaritma fungsi likelihoodnya diperoleh:

$$\ln L = \sum_{1}^{n} Y_i \ln F(\beta' X_i) + \sum_{1}^{n} (1 - Y_i) \ln \left(1 - F(\beta' X_i)\right) \tag{2.20}$$

Kemudian untuk mendapatkan nilai yang maksimum maka turunan persamaan (2.20) terhadap β disamadengankan dengan nol, sehingga dihasilkan:

$$\frac{\partial \ln L}{\partial \beta} = \sum_{1}^{n} \frac{Y_i\, f\left(\beta' X_i\right) X_i}{F\left(\beta' X_i\right)} + \frac{-(1 - Y_i)\, f\left(\beta' X_i\right) X_i}{1 - F\left(\beta' X_i\right)}$$

$$= \sum_{1}^{n} \frac{Y_i\, f\left(\beta' X_i\right) X_i \left(1 - F\left(\beta' X_i\right)\right) - (1 - Y_i)\, F\left(\beta' X_i\right) f\left(\beta' X_i\right) X_i}{F\left(\beta' X_i\right)\left(1 - F\left(\beta' X_i\right)\right)}$$

$$= \sum_{1}^{n} \frac{Y_i\, f\left(\beta' X_i\right) X_i - Y_i\, f\left(\beta' X_i\right) F\left(\beta' X_i\right) - F\left(\beta' X_i\right) f\left(\beta' X_i\right) X_i + Y_i\, f\left(\beta' X_i\right) F\left(\beta' X_i\right)}{F\left(\beta' X_i\right)\left(1 - F\left(\beta' X_i\right)\right)}$$

$$= \sum_{1}^{n} \frac{Y_i\, f\left(\beta' X_i\right) X_i - F\left(\beta' X_i\right)\, f\left(\beta' X_i\right) X_i}{F\left(\beta' X_i\right)\left(1 - F\left(\beta' X_i\right)\right)}$$

$$= \sum_{1}^{n} \frac{\left(Y_i - F(\beta' X_i)\right)\, f(\beta' X_i) X_i}{F\,(\beta' X_i)\left(1 - F(\beta' X_i)\right)} = 0 \qquad (2.21)$$

Karena model probit mengandaikan ε_i berdistribusi normal $N(0,\sigma^2)$, maka fungsi likelihoodnya yang telah dilogaritmakan (*log-likelihood*) menjadi:

$$\ln L = \sum_{Y_i=0} \ln\left(1 - \Phi\left(\beta' X\right)\right) + \sum_{Y_i=1} \ln \Phi\left(\beta' X_i\right) \qquad (2.22)$$

dimana $\Phi(\beta' X_i)$ adalah fungsi distribusi dari peubah acak yang berdistribusi normal. Turunan pertama dalam memaksimumkan L adalah:

$$\frac{\partial \ln L}{\partial \beta} = \sum_{Y_i=0} \frac{-\emptyset_i}{1 - \Phi_i} X_i + \sum_{Y_i=1} \frac{\emptyset_i}{1 - \Phi_i} X_i$$

$$= \sum_{Y_i=0} \lambda_i^0\, X_i + \sum_{Y_i=1} \lambda_i^1\, X_i$$

$$= \sum_{1}^{n} \left[\frac{q_i \emptyset(q_i \beta' X_i)}{\Phi(q_i \beta' X_i)}\right] X_i, \text{ dengan } q_i = 2Y_i - 1$$

$$= \sum_{1}^{n} \lambda_i\, X_i = 0 \qquad (2.23)$$

I. Distribusi Normal Bivariat

Menurut Johnson dan Wichern (2002:151) distribusi normal bivariat merupakan bentuk pengembangan dari distribusi normal univariat. Adapun bentuk distribusi normal univariat dengan mean μ dan varian σ^2 adalah

$$f\,(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-[(x-\mu)/\sigma]^2/2} \;, -\infty < x < \infty \qquad (2.24)$$

Misalkan akan dilakukan evaluasi parameter distribusi normal bivariat $\mu_1 = E(X_1)$, $\mu_2 = E(X_2)$, $\sigma_{11} = \text{Var}(X_1)$, $\sigma_{22} = \text{Var}(X_2)$, dan $\rho_{12} = \sigma_{12} / (\sqrt{\sigma_{11}}\sqrt{\sigma_{22}}) = \text{Corr}(X_1,X_2)$. Dengan melakukan penginversan matrik kovarian

$$\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{bmatrix}$$

diperoleh

$$\Sigma^{-1} = \frac{1}{\sigma_{11}\sigma_{22} - \sigma_{12}^2} \begin{bmatrix} \sigma_{22} & -\sigma_{12} \\ -\sigma_{12} & \sigma_{11} \end{bmatrix}$$

Dengan koefisien korelasi $\sigma_{12} = \rho_{12}\sqrt{\sigma_{11}}\sqrt{\sigma_{22}}$ maka diperoleh $\sigma_{11}\sigma_{22} - \sigma_{12}^2 = \sigma_{11}\sigma_{22}(1 - \rho_{12}^2)$ dan jarak kuadratnya menjadi

$$(x-\mu)'\Sigma^{-1}(x-\mu)$$

$$= [x_1 - \mu_1, x_2 - \mu_2] \frac{1}{\sigma_{11}\sigma_{22} - \sigma_{12}^2} \begin{bmatrix} \sigma_{22} & -\rho_{12}\sqrt{\sigma_{11}}\sqrt{\sigma_{22}} \\ -\rho_{12}\sqrt{\sigma_{11}}\sqrt{\sigma_{22}} & \sigma_{11} \end{bmatrix} \begin{bmatrix} x_1 - \mu_1 \\ x_2 - \mu_2 \end{bmatrix}$$

$$= \frac{\sigma_{22}(x_1 - \mu_1)^2 + \sigma_{11}(x_2 - \mu_2)^2 - 2\rho_{12}\sqrt{\sigma_{11}}\sqrt{\sigma_{22}}(x_1 - \mu_1)(x_2 - \mu_2)}{\sigma_{11}\sigma_{22}(1 - \rho_{12}^2)}$$

$$= \frac{1}{1 - \rho_{12}^2}\left[\left(\frac{x_1 - \mu_1}{\sqrt{\sigma_{11}}}\right)^2 + \left(\frac{x_2 - \mu_2}{\sqrt{\sigma_{22}}}\right)^2 - 2\rho_{12}\left(\frac{x_1 - \mu_1}{\sqrt{\sigma_{11}}}\right)\left(\frac{x_2 - \mu_2}{\sqrt{\sigma_{22}}}\right)\right]$$

Selanjutnya, karena $|\sum| = \sigma_{11}\sigma_{22} - \sigma_{12}^2(1 - \rho_{12}^2)$ maka $\sum^{-1}$ dan $|\sum|$ dapat disubstitusikan kedalam persamaan

$$f(x) = \frac{1}{(2\pi)^{p/2}|\sum|^{1/2}} e^{-(x-\mu)'\Sigma^{-1}(x-\mu)/2} \qquad (2.25)$$

untuk mendapatkan bentuk distribusi normal bivariat beserta parameternya μ_1, μ_2, σ_{11}, σ_{22}, dan ρ_{12}. Berikut adalah bentuk persamaan distribusi normal bivariatnya.

$$f(x_1, x_2) = \frac{1}{2\pi\sqrt{\sigma_{11}\sigma_{22}(1-\rho_{12}^2)}} \exp\left\{-\frac{1}{2(1-\rho_{12}^2)}\left[\left(\frac{x_1-\mu_1}{\sqrt{\sigma_{11}}}\right)^2 + \left(\frac{x_2-\mu_2}{\sqrt{\sigma_{22}}}\right)^2 - 2\rho_{12}\left(\frac{x_1-\mu_1}{\sqrt{\sigma_{11}}}\right)\left(\frac{x_2-\mu_2}{\sqrt{\sigma_{22}}}\right)\right]\right\} \tag{2.26}$$

J. Matriks Hessian

Matriks Hessian adalah matriks persegi dari turunan parsial orde kedua (Agresti, 1990). Misal didefinisikan fungsi riil *f* sebagai berikut: $f\,(x_1, x_2, \ldots, x_n)$. Jika turunan parsial orde kedua untuk semua *f* terdefinisi, maka matriks Hessian dari fungsi *f* adalah:

$$\mathbf{H}(x) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$$

K. Metode Newton-Raphson

Metode Newton-Raphson adalah suatu metode untuk menyelesaikan sistem persamaan yang tidak linier (Agresti, 1990). Metode Newton-Raphson dapat dikembangkan dari perluasan deret Taylor, yang dapat dinyatakan sebagai:

$$f(x) = f(x_n) + (x - x_n)f'(x_n) + \frac{(x - x_n)^2}{2!} f''(x_n) + \ldots \tag{2.27}$$

untuk $n = 0, 1, 2, \ldots$

Suku-suku orde kedua dari perluasan deret Taylor disekitar x_n adalah:

$$f(x) = f(x_n) + (x - x_n)f'(x_n) + \frac{(x - x_n)^2}{2!} f''(x_n) \tag{2.28}$$

Jika x terdiri dari $x_1, x_2, \ldots, x_m$ dan $f\,(x_1, x_2, \ldots, x_m)$ dapat ditulis $\boldsymbol{f(x)}$,

$$f'(\boldsymbol{x}) = \boldsymbol{T}(\boldsymbol{x}) = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix} \text{ serta } \boldsymbol{f''}(\boldsymbol{x}) = \mathbf{H}(x) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_m} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_m \partial x_1} & \frac{\partial^2 f}{\partial x_m \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_m^2} \end{bmatrix}$$

Maka persamaan (2.28) dapat ditulis dengan:

$$\boldsymbol{f}(\boldsymbol{x}) = f(x_n) + (x - x_n)T(x_n) + \frac{(x - x_n)^2}{2}\,\mathbf{H}(x_n) \tag{2.29}$$

Turunan dari persamaan (2.29) terhadap x adalah:

$$\boldsymbol{T}(x) = \boldsymbol{T}(x_n) + (x - x_n)\,\mathbf{H}(x_n) \tag{2.30}$$

Jika $\boldsymbol{T}(x) = 0$, maka akan diperoleh :

$$\boldsymbol{T}(x_n) + (x - x_n)\,\mathbf{H}(x_n) = 0 \tag{2.31}$$

Pendekatan yang baik dari x_n adalah x_{n+1}, maka persamaan (2.31) dapat ditulis:

$$\boldsymbol{T}(x_n) + (x_{n+1} - x_n)\,\mathbf{H}(x_n) = 0 \tag{2.32}$$

Dengan menyelesaikan persamaan (2.32), maka dapat diperoleh suatu iterasi berikut:

$$\boldsymbol{T}(x_n) + (x_{n+1} - x_n)\,\mathbf{H}(x_n) = 0$$

$$(x_{n+1} - x_n)\,\mathbf{H}(x_n) = -\,\boldsymbol{T}(x_n)$$

$$(x_{n+1} - x_n) = -\,\boldsymbol{T}(x_n)\,\mathbf{H}(x_n)^{-1}$$

$$x_{n+1} = x_n - \boldsymbol{T}(x_n)\,\mathbf{H}(x_n)^{-1} \tag{2.33}$$

L. Lagrange Multiplier

Lagrange Multiplier digunakan untuk mengetahui galat pada dua persamaan apakah keduanya pada masing-masing variabel tak bebasnya secara signifikan saling berkorelasi atau tidak (Agresti, 2007: 10).

Adapun langkah-langkah pengujian untuk mengetahui ada tidaknya korelasi antara galat masing-masing model dengan menggunakan uji Lagrange Multiplier adalah:

1. Perumusan Hipotesis

   $H_0 : \rho = 0$

   $H_1 : \rho \neq 0$

2. Besaran yang diperlukan

   Menghitung $g = \sum_{i=1}^{n} q_{i1}q_{i2} \frac{\emptyset(w_{i1})\ \emptyset(w_{i2})}{\Phi(w_{i1})\ \Phi(w_{i2})}$ dan

   $$h = \sum_{i=1}^{n} q_{i1}q_{i2} \frac{[\emptyset(w_{i1})\ \emptyset(w_{i2})]^2}{\Phi(w_{i1})\Phi(-w_{i1})\ \Phi(w_{i2})\Phi(-w_{i2})}$$

3. Statistik Uji

   $$LM = \frac{g^2}{h}$$

4. Kriteria Pengujian

   Dengan mengambil taraf signifikansi α, maka H_0 ditolak jika *p-value* < α.

5. Kesimpulan

   Penafsiran H_0 ditolak memberi arti bahwa korelasi antara galat masing-masing model adalah tidak sama dengan nol atau dengan kata lain bahwa kedua model persamaan secara signifikan saling berkorelasi satu sama lain.

M. Uji Perbandingan Likelihood

Menurut Agresti (2007: 10) uji hipotesis bagi koefisien regresi secara simultan dilakukan dengan maksud untuk mengetahui apakah variabel-variabel bebas yang digunakan dalam model secara simultan mempunyai pengaruh terhadap variabel yang ingin dijelaskan atau tidak. Pada model regresi probit bivariat digunakan uji perbandingan likelihood untuk menguji parameter secara simultan.

Langkah-langkah yang perlu dilakukan dalam pengujian signifikansi parameter secara simultan dengan menggunakan uji perbandingan likelihood adalah sebagai berikut:

1. Perumusan Hipotesis

   H_0 : $\beta_{j1} = ... = \beta_{jp} = 0$, untuk $j = 1, 2$

   H_1 : sekurang-kurangnya terdapat satu $\beta_{jk} \neq 0$, untuk $j = 1, 2$, $k = 1, ..., p$

2. Besaran yang diperlukan

   Menghitung $-2\log\left(\frac{\text{likelihood tanpa variabel bebas}}{\text{likelihood dengan variabel bebas}}\right)$ dengan bantuan *software* Stata versi 10.

3. Statistik Uji

   $$\chi^2_{hitung} = -2\log\left(\frac{\text{likelihood tanpa variabel bebas}}{\text{likelihood dengan variabel bebas}}\right)$$

4. Kriteria Pengujian

   Dengan taraf signifikansi $\alpha = 0.05$, maka H_0 ditolak jika $\chi^2_{hitung} > \chi^2_{(1-\alpha);p}$

5. Kesimpulan

   Penafsiran dari H_0 diterima atau di tolak.

N. Uji Wald

Uji Wald, menurut Agresti (2007: 11), digunakan untuk menguji signifikansi masing-masing parameter. Statistik uji Wald dihitung dengan membagi parameter yang ditaksir oleh galat baku dari parameter yang ditaksir tersebut, yaitu:

$$Z_{jk} = \frac{\hat{\beta}_{jk}}{SE(\hat{\beta}_{jk})} \qquad j = 1, 2 \tag{2.34}$$

dimana $\hat{\beta}_{jk}$ adalah penaksir β_{jk} dan $SE(\hat{\beta}_{jk})$ adalah penaksir galat baku β_{jk}. Adapun langkah-langkah pengujian signifikansi parameter regresi secara parsial dalam uji Wald adalah sebagai berikut:

1. Perumusan Hipotesis

   H_0 : $\beta_{jk} = 0$, untuk $k = 0, 1, ..., p$

   H_1 : $\beta_{jk} \neq 0$, untuk $k = 0, 1, ..., p$

2. Besaran yang diperlukan

   Menghitung $\hat{\beta}_{jk}$dan $SE(\hat{\beta}_{jk})$

3. Statistik Uji

   $$Z_{jk} = \frac{\hat{\beta}_{jk}}{SE(\hat{\beta}_{jk})}$$

4. Kriteria Pengujian

   Dengan mengambil taraf signifikansi $\alpha = 0.05$, maka H_0 diterima jika

   $-z_{\frac{1}{2}(1-\alpha)} < z_{jk} < z_{\frac{1}{2}(1-\alpha)}$

5. Kesimpulan

   Penafsiran H_0 diterima atau ditolak.